

DTIC FILE COPY

1

AD \_\_\_\_\_

AD-A223 055

GRANT NO: DAMD17-88-Z-8035

TITLE: ROUNDTABLE FOR THE DEVELOPMENT OF DRUGS AND VACCINES  
AGAINST ACQUIRED IMMUNO DEFICIENCY SYNDROME (AIDS)

SUBTITLE: Surrogate Endpoints in Evaluating the Effectiveness  
of Drugs Against HIV Infection and AIDS

PRINCIPAL INVESTIGATOR: Robin Weiss, M.D.  
Leah Mazade

CONTRACTING ORGANIZATION: National Academy of Sciences  
2101 Constitution Avenue  
Washington, DC 20418

REPORT DATE: June 1, 1990

TYPE OF REPORT: Final Proceedings

PREPARED FOR: U.S. ARMY MEDICAL RESEARCH AND DEVELOPMENT COMMAND  
Fort Detrick, Frederick, Maryland 21701-5012

DISTRIBUTION STATEMENT: Approved for public release;  
distribution unlimited

The findings in this report are not to be construed as an  
official Department of the Army position unless so designated by  
other authorized documents.

DTIC  
S ELECTE  
JUN 25 1990  
6 D

90 06 25 023

## REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified			1b. RESTRICTIVE MARKINGS		
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited		
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE			5. MONITORING ORGANIZATION REPORT NUMBER(S)		
4. PERFORMING ORGANIZATION REPORT NUMBER(S)			7a. NAME OF MONITORING ORGANIZATION		
6a. NAME OF PERFORMING ORGANIZATION National Academy of Sciences		6b. OFFICE SYMBOL (if applicable)	7b. ADDRESS (City, State, and ZIP Code)		
6c. ADDRESS (City, State, and ZIP Code) 2101 Constitution Avenue Washington, DC 20418			9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER Grant No. DAMD17-88-Z-8035		
8a. NAME OF FUNDING/SPONSORING ORGANIZATION U.S. Army Medical Research & Development Command		8b. OFFICE SYMBOL (if applicable)	10. SOURCE OF FUNDING NUMBERS		
8c. ADDRESS (City, State, and ZIP Code) Fort Detrick Frederick, Maryland 21701-5012			PROGRAM ELEMENT NO. --	PROJECT NO. --	TASK NO. --
11. TITLE (Include Security Classification) ROUNDTABLE FOR THE DEVELOPMENT OF DRUGS AND VACCINES AGAINST ACQUIRED IMMUNO DEFICIENCY SYNDROME (AIDS)					
12. PERSONAL AUTHOR(S) Robin Weiss, M.D., Leah Mazade					
13a. TYPE OF REPORT Final Proceedings		13b. TIME COVERED FROM ____ TO ____		14. DATE OF REPORT (Year, Month, Day) 1990 June 1	
15. PAGE COUNT					
16. SUPPLEMENTARY NOTATION Subtitle: Surrogate Endpoints in Evaluating the Effectiveness of Drugs Against HIV Infection and AIDS					
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP	1. AIDS, Antiviral Drugs, Vaccines, Conference, Symposium		
--	--	--	2. Human Immunodeficiency Virus, Infections		
--	--	--	3. Acquired Immune Deficiency Syndrome, Medical Research		
19. ABSTRACT (Continue on reverse if necessary and identify by block number)  This conference summary was prepared by the Institute of Medicine's Roundtable for the Development of Drugs and Vaccines Against AIDS, chaired by Harold Ginsberg and Sheldon Wolff and directed by Robin Weiss and Richard Berzon. The document reports major themes of the conference discussions; these themes, however, do not represent policy statements by the Institute of Medicine.					
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION Unclassified		
22a. NAME OF RESPONSIBLE INDIVIDUAL Mrs. Virginia M. Miller			22b. TELEPHONE (Include Area Code) (301) 663-7325		22c. OFFICE SYMBOL SGRD-RMI-S

# **Institute of Medicine**

---

## **Surrogate Endpoints in Evaluating the Effectiveness of Drugs Against HIV Infection and AIDS**

**September 11-12, 1989  
Conference Summary**

---



# Surrogate Endpoints in Evaluating the Effectiveness of Drugs Against HIV Infection and AIDS

September 11–12, 1989  
Conference Summary

*Robin Weiss and Leah Mazade, editors*

The Roundtable for the Development of Drugs  
and Vaccines Against AIDS

Institute of Medicine

National Academy Press  
Washington, D.C. 1990

This conference summary was prepared by the Institute of Medicine's Roundtable for the Development of Drugs and Vaccines Against AIDS, chaired by Harold Ginsberg and Sheldon Wolff and directed by Robin Weiss and Richard Berzon. The document reports major themes of the conference discussions; these themes, however, do not represent policy statements by the Institute of Medicine.

This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The Institute of Medicine was chartered in 1970 by the National Academy of Sciences to enlist distinguished members of appropriate professions in the examination of policy matters pertaining to the health of the public. In this, the Institute acts under both the Academy's 1863 Congressional charter responsibility to be an advisor to the federal government, and its own initiative in identifying issues of medical care, research, and education.

The Roundtable is supported by the American Foundation for AIDS Research, the Pharmaceutical Manufacturers Association, the U.S. Army, the U.S. Public Health Service, and the Department of Veterans Affairs.

AIDS Activities  
Institute of Medicine  
2101 Constitution Avenue, N.W.  
Washington, D.C. 20418  
(202) 334-2453

Publication Number IOM-90-001

# INSTITUTE OF MEDICINE

## Roundtable for the Development of Drugs and Vaccines Against AIDS

**Harold S. Ginsberg, Cochair**, Eugene Higgins Professor of Medicine and Microbiology, Emeritus, Department of Medicine, College of Physicians & Surgeons, Columbia University, New York

**Sheldon M. Wolff, Cochair**, Physician-in-Chief, New England Medical Center, and Endicott Professor and Chairman, Department of Medicine, Tufts University School of Medicine, Boston

**David W. Barry**, Vice President of Research, Development, and Medical Affairs, The Wellcome Research Laboratories, Burroughs Wellcome Co., Research Triangle Park, North Carolina

**Donald S. Burke**, Colonel, Medical Corps, U.S. Army, and Director, Division of Retrovirology, Walter Reed Army Institute of Research, Rockville, Maryland

**Bruce A. Chabner**, Director, Division of Cancer Treatment, National Cancer Institute, National Institutes of Health, Bethesda, Maryland

**Max D. Cooper**, Professor of Pediatrics, Department of Pediatrics, University of Alabama Medical Center, Birmingham

**Martin Delaney**, Co-Executive Director and President, Project Inform, San Francisco

**Daniel Deykin**, Chief, Cooperative Studies Program, Veterans Administration Medical Center, Boston

**Anthony S. Fauci**, Associate Director, Office of AIDS Research, and Director, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland

**Gerald Friedland**, Professor of Medicine, Department of Medicine, Albert Einstein College of Medicine, Montefiore Medical Center, The Bronx

**L. Patrick Gage**, Senior Vice President, Scientific Affairs, Genetics Institute, Inc., Cambridge, Massachusetts

**Robert C. Gallo**, Chief, Laboratory of Tumor Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland

**Peter Barton Hutt**, Partner, Covington & Burling, Washington, D.C.

**Jay C. Lipner**, Partner, Silverstein Langer Lipner & Newburgh, New York

**David W. Martin, Jr.**, Senior Vice President, Research and Development, Genentech, Inc., South San Francisco

**Edward M. Scolnick**, President, Merck Sharp & Dohme Research Laboratories, Merck and Company, Rahway, New Jersey

**Catherine M. Wilfert**, Professor of Pediatrics and Microbiology, Department of Pediatrics, Division of Infectious Diseases, Duke University Medical Center, Durham, North Carolina

**Frank E. Young**, Deputy Assistant Secretary for Health (for Science and the Environment), Department of Health and Human Services, Washington, D.C.

### Institute of Medicine Staff

**Richard Berzon**, Project Officer  
**Leah Mazade**, Consultant  
**Gail Spears**, Administrative Assistant  
**Robin Weiss**, Director



Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution /	
Availability Codes	
Dist	Avail and/or Special
A-1	

## CONTENTS

Conference Program	vi
Conference Participants	viii
Introduction	1
Virologic Markers	3
Immunologic Markers	4
Special Issues in the Maternal and Pediatric Population	7
Endpoints, Trial Design, and Regulation	10
Summary	12

**SURROGATE ENDPOINTS IN EVALUATING THE EFFECTIVENESS  
OF DRUGS AGAINST HIV INFECTION AND AIDS**

**September 11-12, 1989**

**PROGRAM**

**Monday, September 11**

**8:00    *Welcome and Opening Remarks***

- Samuel O. Thier
- Harold Ginsberg
- Sheldon Wolff

**8:15    *Surrogate Endpoints in Clinical Drug Trials***

- James Bilstad

**8:30    *Surrogate Endpoints in AIDS Drug Trials***

- Paul Volberding

**8:45    *Laboratory Predictors of Progression to AIDS: The Epidemiological Data***

- Andrew Moss

**9:15    *Virologic Markers: p24 Core Antigen, Quantitative Plasma Cultures, and Polymerase Chain Reaction***

***Moderator:*** Thomas Merigan

**Panelists**

Henry Balfour  
Douglas Richman  
Robert Coombs  
Stephen Wolinsky

**Discussants**

Paul Beninger  
Lawrence Corey  
Blaine Hollinger

**11:00    Break**

**11:15    *Immunologic Markers: CD4, CD8, Beta-2 Microglobulin, Neopterin, and Soluble Interleukin-2 Receptors***

***Moderator:*** Anthony Fauci

**Panelists**

John Phair  
Margaret Fischl  
Mark Jacobson

**Discussants**

John Fahey  
Janis Giorgi  
Gerald Quinnan  
Fred Valentine

**1:00    Lunch Break**

**2:00    *Special Issues in the Maternal and Pediatric Populations***

***Moderator:*** Stanley Plotkin

**Panelists**

Gwendolyn Scott  
Catherine Wilfert  
Karina Butler

**Discussants**

Edward Connor  
John Modlin  
Paul Parkman



3:30 Break

3:45 *Endpoints, Trial Design, and Regulation*

*Moderator:* Frank Young

**Panelists**

Ellen Cooper  
Martin Delaney  
Sandra Lehrman  
Daniel Hoth

**Discussants**

Peter Barton Hutt  
Stephen Lagakos

5:30 Adjourn

**Tuesday, September 12**

8:30 *Panel Summaries*

- Thomas Merigan
- Anthony Fauci
- Stanley Plotkin
- Frank Young

9:15 *WORKSHOP: Which Surrogate Endpoints are Ready for Use in Evaluating AIDS Clinical Drug Trials?*

*Moderator:* Martin Hirsch

**Discussants**

Samuel Broder  
Ellen Cooper  
Anthony Fauci  
Gerald Friedland  
Clifford Lane

**Discussants**

Jay Lipner  
Thomas Merigan  
Stanley Plotkin  
Stephen Sherwin  
Frank Young

10:30 Break

11:00 *Discussion*

12:00 *Summation*

- Sheldon Wolff

12:30 Adjourn

## CONFERENCE PARTICIPANTS

- Henry H. Balfour, Jr.*, Department of Laboratory Medicine and Pathology and Department of Pediatrics, University of Minnesota Medical School, Minneapolis
- Paul R. Beninger*, Division of Antiviral Drug Products, Center for Drug Evaluation and Research, Food and Drug Administration, Bethesda, Maryland
- James M. Bilstad*, Office of Drug Evaluation II, Center for Drug Evaluation and Research, Food and Drug Administration, Bethesda, Maryland
- Samuel Broder*, National Cancer Institute, National Institutes of Health, Bethesda, Maryland
- Karina M. Butler*, Infectious Disease Section, Pediatric Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland
- Edward Connor*, Division of Pediatrics, Allergy, Immunology and Infectious Diseases, Children's Hospital of New Jersey, Newark
- Robert W. Coombs*, Departments of Medicine and Laboratory Medicine, University of Washington, Seattle
- Ellen C. Cooper*, Division of Antiviral Drug Products, Center for Drug Evaluation and Research, Food and Drug Administration, Bethesda, Maryland
- Lawrence Corey*, Virology Division, University of Washington, Children's Hospital, Seattle
- Martin Delaney*, Project Inform, San Francisco
- John L. Fahey*, Center for Interdisciplinary Research in Immunology and Disease, University of California, Los Angeles
- Anthony S. Fauci*, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland
- Margaret Fischl*, Department of Medicine, University of Miami School of Medicine
- Gerald Friedland*, Department of Medicine, Albert Einstein College of Medicine, The Bronx
- Harold S. Ginsberg*, Department of Medicine, College of Physicians and Surgeons, Columbia University, New York
- Janis V. Giorgi*, Department of Microbiology and Immunology, School of Medicine, University of California, Los Angeles
- Martin S. Hirsch*, Department of Virology, Massachusetts General Hospital, Boston
- F. Blaine Hollinger*, GI Immunodiagnostic and AIDS Virus Laboratories, Baylor College of Medicine, Houston
- Daniel F. Hoth*, Division of the Acquired Immunodeficiency Syndrome, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland
- Peter Barton Hut*, Covington & Burling, Washington, D.C.
- Mark A. Jacobson*, Division of Infectious Disease and Clinical Pharmacology, Department of Medicine, University of California, San Francisco
- Stephen Lagakos*, Department of Biostatistics, Harvard University School of Public Health, Boston
- H. Clifford Lane*, Laboratory of Immunoregulation, National Institute of Allergy and Infectious Diseases of the National Institutes of Health, Bethesda, Maryland
- Sandra Nusinoff Lehrman*, Department of Antimicrobial Therapy, Burroughs Wellcome Company, Research Triangle Park, North Carolina
- Jay C. Lipner*, Silverstein Langer Lipner & Newburgh, New York
- Thomas C. Merigan, Jr.*, Center for AIDS Research, Stanford University School of Medicine
- John F. Modlin*, Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, Maryland
- Andrew R. Moss*, Department of Epidemiology and Biostatistics, University of California, San Francisco
- Paul D. Parkman*, Center for Biologics Evaluation and Research, Food and Drug Administration, Bethesda, Maryland
- John Phair*, Department of Medicine, Northwestern University Medical School, Chicago

*Stanley A. Plotkin*, Division of Infectious Diseases, Children's Hospital of Philadelphia  
*Gerald V. Quinnan, Jr.*, Center for Biologics Evaluation and Research, Food and Drug  
Administration, Bethesda, Maryland  
*Douglas D. Richman*, Clinical Virology Section, Laboratory Service, San Diego Veterans  
Administration Medical Center  
*Gwendolyn B. Scott*, Division of Pediatric Immunology and Infectious Diseases, University of Miami  
School of Medicine  
*Stephen A. Sherwin*, Clinical Research, Genentech, Inc., San Francisco  
*Samuel O. Thier*, Institute of Medicine, National Academy of Sciences, Washington, D.C.  
*Fred T. Valentine*, Department of Medicine, New York University School of Medicine  
*Paul Volberding*, Medical Oncology Division, San Francisco General Hospital  
*Catherine M. Wilfert*, Department of Pediatrics, Duke University Medical Center, Durham, North  
Carolina  
*Sheldon M. Wolff*, Department of Medicine, Tufts University School of Medicine, Boston  
*Stephen Wolinsky*, Department of Medicine, Northwestern University School of Medicine, Chicago  
*Frank E. Young*, Food and Drug Administration, Bethesda, Maryland

## INTRODUCTION\*

In the search for effective therapies against infection by the human immunodeficiency virus (HIV), patients and their care givers have increasingly called for a shortening of the testing and approval process that has traditionally brought drugs from development to market. As regulated by the Food and Drug Administration (FDA), this process often takes from 7 to 10 years and involves preclinical testing and a phased series of clinical trials of increasing size and complexity to investigate a new drug's potential toxicity and effectiveness. In the drug approval process, FDA has sought to achieve a balance between requiring definitive, large-scale, carefully monitored clinical trials to determine the safety and efficacy of new drugs and ensuring that patients receive helpful medicines at the earliest possible time. Yet in the case of HIV infection, that balance has proved elusive. The protracted course of the disease and the fact that patients may be infected but remain asymptomatic for a relatively long period compound the difficulties of large-scale trials and could lengthen even more the time required under traditional methods to establish the efficacy of a drug. The pressure for more rapid approval of effective therapies has led to a search for appropriate ways to shorten the time it takes to establish drug efficacy. One such mechanism that has been used by FDA in other drug classes and that may hold promise for HIV infection is the surrogate endpoint.

To define a surrogate endpoint, it is first necessary to define the "true" clinical endpoint for which the surrogate will substitute. A true endpoint, often termed the significant clinical endpoint, is that outcome measure by which therapies are deemed efficacious or ineffectual in clinical trials. In the case of AIDS it is generally death (mortality) or serious disease (morbidity—usually, an opportunistic infection). A surrogate endpoint, then, is a substitute for these later, definitive outcomes that nevertheless correlates with disease progression and shows changes as a result of drug effects.

A surrogate endpoint can involve determining a laboratory measurement or marker (e.g., CD4+ cell counts, levels of p24 antigen) that reflects in some manner the status of the disease process. Clinical conditions (e.g., changes in weight, fever, neuropsychiatric functions) may also be measured, and, to the extent that they predict more severe disease or survival, some researchers consider them surrogate markers as well. (An important difference between laboratory and clinical measures is that clinical measures may have independent significance for patient well-being as well as usefulness for predicting later disease.) Although there is no universally accepted definition of a surrogate endpoint, there is a distinction to be made between a surrogate endpoint and a laboratory marker of disease progression. A particular laboratory measurement may be a good predictor of the course of the disease, but it may not necessarily be an appropriate surrogate endpoint for drug efficacy trials. There is no guarantee that a marker that predicts progression of disease in an untreated patient will respond in a reverse fashion to therapeutic intervention.

One of the primary benefits of establishing a surrogate endpoint for drug trials is that any effects of a particular therapy can be seen earlier in the course of clinical trials before patients progress to a more severe disease state. This early determination of effects allows for quicker approval and more widespread availability of an effective treatment. Surrogate endpoints also have other advantages. In some cases, they may be easier to measure than the significant clinical endpoint; in addition, drug development costs may be lower because long, complex drug efficacy trials with large populations are shortened or simplified. Surrogate endpoints may also be used at different stages of the clinical trials process. They can serve as the primary endpoint in Phase I trials of drug activity or as secondary endpoints in efficacy trials to support mortality and morbidity findings. In addition, they can be used as stratification criteria to divide a population into groups for consideration of drug effects on different levels of the marker being used.

In addition to benefits, however, there are risks to be considered in replacing clinical endpoints with surrogates that rely on measurement of laboratory markers. A drug that was actually useful

---

\*This section is based on material presented by James Bilstad, Paul Volberding, Andrew Moss, and Lawrence Corey.

might be rejected (or considered to be of little interest) simply because it showed no effect on the particular marker chosen as a surrogate. The reverse also applies: a drug that actually produced little clinical benefit might nevertheless be considered efficacious because it affected the marker favorably. These considerations argue for great care in choosing and validating surrogate endpoints.

Another factor to be considered is that good markers have characteristics that may vary from population to population. For example, it appears that beta-2 microglobulin is not as strongly predictive of disease progression among intravenous drug users as it is among homosexual men. Also, p24 antigenemia shows population variability: the marker appears to be found less frequently among blacks than among whites. This type of variability increases the difficulties already inherent in the task of identifying an appropriate surrogate for the true clinical endpoints of AIDS drug efficacy trials.

An appropriate surrogate endpoint must satisfy a number of criteria:

1. The marker must be biologically plausible; that is, it must be consistent with what is known about the pathophysiology and pathogenesis of the disease. If a marker is central to the pathogenesis of the disease, then regression of the marker (i.e., the reversal or stabilization of an adverse trend) should occur with therapies that have different mechanisms of action.
2. The marker must be present or abnormal in a large percentage of people who have the disease. Preferably, it would be present in all persons with HIV infection.
3. The marker must be a good predictor of disease progression and should correlate closely with the significant clinical endpoint.
4. There should be a correlation between the quantitative aspect of the marker and the progression of the infection, i.e., the more severe the infection, the more deviant the marker from normal.
5. The regression of the marker should be significantly associated with clinical improvement (i.e., those with the greatest improvement in the marker should also show the greatest clinical effects). Conversely, the lack of regression of the marker should be commonly associated with a lack of clinical improvement.
6. The incidence of regression or improvement in the laboratory marker should be significantly greater in treated than in untreated patients.

Whether a marker holds up as a surrogate for the clinical endpoint in a drug trial is eventually determined by validating the marker through well-designed, carefully conducted, prospective clinical studies that establish correlations between surrogate and clinical outcomes. Such validation means that the drug must be studied for its effect on both the clinical endpoint and on the surrogate marker. The current trend toward informal or truncated clinical trials is of some concern as it may damage the potential for obtaining the required correlations. Once a drug is approved with well-designed studies that validate the surrogate endpoint, it may be easier to demonstrate the effectiveness of subsequent drugs of the same class on the basis of the surrogate endpoint. The weight given to any particular marker depends on how well it correlates with the clinical endpoint and the degree to which it has been validated. Some markers may be useful for determining drug activity or dose ranges in early clinical studies, but they may not be useful as primary efficacy endpoints for approval of the drug for marketing. In addition, a marker may be validated for use as a surrogate endpoint for one stage of the disease but not for all stages. For instance, a marker may be valid and valuable in asymptomatic infection but lose its value in later stages of disease. Some of the drug classes in which FDA has accepted surrogate endpoints as a basis for approval include antihypertensive drugs, drugs to treat osteoporosis, vaccines, lipid-altering drugs, antiarrhythmic drugs, and alpha-1 proteinase inhibitors. With HIV infection, however, the question is whether a drug should be approved now on the basis of a surrogate endpoint that may or may not be validated by future studies.

Sometimes a candidate surrogate endpoint is initially identified through epidemiological studies in which it is found to be a strong predictor of disease progression. During clinical trials (where the bulk of experience so far is with nucleoside analogues), reproducible, measurable clinical benefits

(e.g., improved neurological status) other than the primary endpoints of interest may also be observed, offering another way to measure drug efficaciousness. Predictors of progression to AIDS gleaned from epidemiological data include both virologic markers (e.g., p24 antigen levels, plasma viremia) and immunologic markers (e.g., CD4+ cell counts, beta-2 microglobulin, neopterin). The potential use of these various markers as surrogate endpoints for drug efficacy trials is discussed in the following sections.

### VIROLOGIC MARKERS\*

In the search for surrogate endpoints of drug efficacy against HIV, researchers have quite naturally looked to the virus itself for a virologic marker of disease progression. They have investigated viral burden using a number of different techniques and considering several specific elements of the virus. Particular attention has centered on p24 antigen, the major viral core protein of HIV. As a result the measurement of p24 has been included in clinical trials to determine whether the administration of a particular agent appears to affect levels of the antigen. Indeed, such measurement was used to advantage in the Phase II trial of AZT. The p24 antigen is relatively easy to detect and quantify, and measurement methods have been standardized and validated largely as a result of the work of the Virology Committee of the National Institutes of Health (NIH) AIDS Clinical Trials Group (ACTG). Measurable levels of the antigen have been shown to correlate with stages of disease and disease progression. For example, one study showed more symptomatic disease among hemophiliacs with detectable levels of p24 antigen than among those who were not antigenemic. In addition, p24 levels have been shown to decrease with the administration of antiretroviral drugs—to the point that in some studies a dose-response relationship has been seen. Furthermore, several drugs that have failed to affect serum or plasma levels of p24 antigen have shown few indications of efficacy by other laboratory or clinical measures.

Despite these advantages, the use of p24 levels as a surrogate endpoint in drug trials has several major drawbacks. First, the antigen is not present in all HIV-infected persons; in fact, the percentage of individuals who show p24 antigenemia at any stage of infection, and particularly in asymptomatic infection, is relatively low. Any clinical trial enrollment limited to p24 antigenemic persons would be severely constrained in its population. Moreover, although a reduction of p24 levels following the administration of AZT has been correlated with clinical response, this remains only an association. There is no guarantee that an antigen response to therapy will consistently predict a clinical response or that failure to decrease the level of p24 antigen correlates with a poor prognosis.

From a biological perspective, it is premature to accept p24 antigen as a surrogate endpoint for drug efficacy. For one thing, it is not known which cells *in vivo* produce the antigen or what the reduction of p24 after therapy means at a cellular level. Consequently, although many researchers display some confidence that the reduction of p24 with the use of nucleoside analogues (e.g., AZT, ddI, ddC) reflects antiretroviral activity, they are unable to explain why this reduction occurs. Also unknown is the route by which the antigen enters the circulation or why many HIV-infected individuals develop AIDS—and die—without ever generating measurable antigen levels. When researchers measure p24 antigen, they are really measuring only one component of a dynamic system comprising antigen, antibody, and immune complexes (which may interfere with measuring free antigen). So far the majority of attention has been focused on free antigen. Yet some investigators have proposed that the drop in antibody seen in later stages of HIV infection may actually be a better predictor of disease progression than antigen levels, especially when used in combination with other indicators. A further limitation to the use of p24 antigen as a surrogate

---

\* This section is based on material presented by Thomas Merigan, Henry Balfour, Douglas Richman, Robert Coombs, Stephen Wolinsky, Paul Beninger, Lawrence Corey, and Blaine Hollinger.

endpoint is that criteria for what constitutes a significant decline in p24 levels in clinical trials have not been established. This limitation, as well as the drawbacks noted earlier, must be addressed before p24 antigen can serve as a major surrogate marker of drug efficacy.

Virus cultured from peripheral blood mononuclear cells, although often considered by virologists to be a "gold standard" against which to measure other markers, has not proved particularly useful in the case of HIV infection. Current measurement techniques are so sensitive that the virus can be reliably isolated from the great majority of antibody-positive individuals, regardless of the stage of their disease. Once the culture is positive, it remains so, and there has been no evidence to date that AZT reverses this result. The lack of change with therapy in the ability to culture cell-associated virus suggests that, despite therapeutic progress, either drugs with greater antiviral effects are needed or culturing techniques amplify latent virus that antivirals in theory cannot eradicate.

Unlike p24 antigen, plasma viremia (as measured by virus cultured from plasma) is highly prevalent in persons with AIDS. In recent cross-sectional studies that quantified plasma viremia by serial dilution, plasma viremia rose from approximately 20 percent in asymptomatic subjects to 90 percent in AIDS patients. Moreover, a prospective study by the same investigators showed that the appearance of plasma viremia was associated with progression to symptomatic disease and declines in CD4+ cell counts. In preliminary data presented at the conference, oral zidovudine therapy did not significantly reduce the titer or frequency of plasma viremia, although there was a downward trend with therapy. The same patients displayed a significant decline in antigen level with therapy. Because plasma viremia is a useful marker of disease progression, delay in the progression of plasma viremia may become a valuable means of evaluating response to antiviral therapy early in infection.\*

Another area of exploration that appears promising is the use of polymerase chain reaction (PCR) to establish and measure viral burden. PCR is a gene amplification technique that detects proviral sequences in the peripheral blood mononuclear cells of HIV-infected persons. Thus far, PCR has been used mainly to establish the presence of infection. Its use as a marker of drug efficacy is more problematic. For example, any alterations in the efficiency of the process may substantially affect the reliability of quantification. In addition, the technique is not yet widely available, and laboratories have yet to standardize procedures and establish reliability. PCR has been used to study a few patients before and after AZT therapy, and the studies suggested decreases in the amount of detectable provirus after AZT administration. Such efforts, although in their infancy, point to the value of continuing investigation of PCR and suggest it may be a technique that holds promise for the future.

### IMMUNOLOGIC MARKERS\*\*

The common denominator of HIV infection is a progressive diminution of immune function. It is thus appropriate to consider a variety of immunological variables as surrogate endpoints in efficacy trials for drugs against AIDS and HIV. Data have been gathered on several markers of immune system activation, such as beta-2 microglobulin, neopterin, and soluble interleukin-2 receptors. Yet

---

\* As evidence of how quickly the field is developing, soon after this conference, two studies were published that amplify these conclusions. In one (Ho, D.D., T. Moudgil, and M. Alam, "Quantitation of human immunodeficiency virus type 1 in the blood of infected persons," *New England Journal of Medicine*, vol. 321, no. 24, pp. 1621-1625 [1989]), both plasma and peripheral blood mononuclear cell end-point-dilution cultures were found to vary in titers with stage of disease. In Ho's study, the HIV-1 titers of seven patients with AIDS or ARC (AIDS-related complex) treated with AZT declined significantly in plasma but not in peripheral blood mononuclear cells. Coombs and colleagues (Coombs, R.W., A.C. Collier, J.-P. Allain, B. Nikora, M. Leuther, G.F. Gjeret, and L. Corey, "Plasma viremia in human immunodeficiency virus infection," *New England Journal of Medicine*, vol. 321, no. 24, pp. 1626-1331 [1989]) elaborated on the results presented here. Both studies raise the possibility of a role for plasma viremia as a surrogate marker of drug efficacy.

\*\*This section is based on material presented by Anthony Fauci, John Phair, Margaret Fischl, Mark Jacobson, Janis Giorgi, Gerald Quinnan, and Fred Valentine.

the marker that continues to hold the most promise as a surrogate endpoint is the number of CD4+ cells. This immunological indicator reflects an essential component of the human immune system that is affected directly by HIV and that shows a clear relationship to disease pathogenesis. In addition, studies have increasingly shown the value of this measure as a predictor of clinical outcome.

The CD4+ cell is a T lymphocyte (its precursors lie in the thymus) that expresses the cell surface marker CD4 to which HIV binds. Thus, it is the primary target for HIV. CD4+ cells are also essential in the coordination of many critical immune system functions. Many of the immunological abnormalities in individuals with HIV infection are secondary to the lack of an appropriate inductive signal from the CD4+ cell. Measurement of CD4+ cell counts has become embedded in clinical management and clinical trials, and data on this marker over a relatively long period are now available. They show that there is an unequivocal association between low levels of CD4+ cells and measurable deleterious events (i.e., initial AIDS-defining events such as *Pneumocystis carinii* pneumonia [PCP] or wasting, as well as their recurrence) over the course of HIV infection. The reverse also appears to hold: the higher the CD4+ cell count, the better the prognosis.

Although more remains to be learned about the patterns and mechanisms of CD4+ cell depletion in HIV disease, studies have identified several periods during which CD4+ cell patterns change. At HIV seroconversion, for example, a rapid fall occurs in CD4+ cell levels. During the period of seropositivity, cell levels often remain stable, but they generally fall rapidly just prior to the development of AIDS. During AIDS, depending on therapy, cell levels usually continue to fall to very low levels. As noted earlier, a strong correlation is seen between CD4+ cell numbers and deleterious events.

One of the studies that has been instrumental in establishing this correlation is the Multicenter AIDS Cohort Study (MACS), an ongoing prospective investigation of the natural history of HIV infection in approximately 5,000 homosexual and bisexual men. Begun in 1984, cumulative data from the MACS indicated that seropositive persons with higher CD4+ counts had a lower risk of developing PCP, whereas those with CD4+ cell counts of less than 200 had a markedly increased risk of developing it. These findings served as the basis for the Public Health Service's recommendation that prophylaxis against PCP should be begun for patients with CD4+ counts of less than 200.

Data are also available on the effects of AZT on CD4+ cells. The National Cancer Institute's (NCI) original Phase I AZT study showed that CD4+ counts increased with administration of the drug, then declined slowly. At the highest doses of the drug, however, CD4+ levels decreased, perhaps as a result of toxicity. In the double blind, placebo-controlled trial of AZT, CD4+ cells increased in the treatment group and then declined slowly, as opposed to the pattern of progressive decline in cell counts seen in the placebo arm of the trial. The preliminary analysis of data from a study investigating progression to AIDS or advanced ARC in individuals with early ARC treated with AZT versus placebo found differences in drug effect between lower and higher CD4+ cell strata. Individuals with CD4+ counts of 200 to 500 treated with AZT showed a mean increase of approximately 50 cells, followed by a slow, progressive decline. Individuals in the higher stratum (CD4+ counts greater than 500) showed only a nominal increase compared to controls, also followed by a slow decline.

An important question that remains unanswered is the meaning of the decline that occurs following the initial increase in CD4+ cells on therapy. Is the decline a reflection of the lymphosuppressive effects of nucleoside analogues, or does it reflect the possibility that the immune system may not be able to reconstitute itself because it has undergone irreparable damage to the thymic precursors of mature CD4+ cells? Further data that have recently become available from a multicenter ACTG study investigating two different doses of AZT in the treatment of AIDS found no difference between the two doses in their effects on CD4+ cells (i.e., CD4+ cells increased early in therapy and then slowly declined). More important, however, these data suggest that the absolute number of CD4+ cells at any one point in time, the latest measurement, is the most important predictor of survival, regardless of how the individual achieved that level (through therapy or by



starting at a different baseline level from which progression or regression occurred). This study also suggested that relatively small shifts in the number of CD4+ cells can be very important if they bring an individual's cell count into a better prognostic category. From these findings comes the hypothesis that a stable CD4+ level indicates that HIV disease is not progressing; therefore, a halt in the decline in the number of CD4+ cells is viewed as meaningful. Such a halt becomes especially important if it prevents an individual from moving to a CD4+ cell count that predicts a deleterious event.

The ability to reliably measure small changes in the number of CD4+ cells is crucial to this marker's use as a predictor of disease progression and potential surrogate endpoint. However, because CD4+ cell absolute values are obtained from three different laboratory measures using three separate instruments, reliable measurement has been a problem. There have been concerted efforts to standardize such measurements and implement quality control in laboratory testing, and a number of laboratories have been certified for such testing. Nevertheless, reliability remains an issue.

In addition to measurement problems, several factors in individual patients may affect CD4+ values. Diurnal variations may occur, or concurrent common viral infections may cause a transient depression of the number of CD4+ cells. The difficulties involved in determining absolute CD4+ cell counts have led some researchers to suggest that more reliable measurements might be obtained by measuring CD4+ cells as a percentage of total lymphocytes (there appears to be less day-to-day fluctuation in measuring the percentage). Another alternative is to use a combination of several measures—for example, cell count, percentage of total lymphocytes, and the CD4+/CD8+ ratio.

Unlike CD4+ cell counts, the other immunological markers noted earlier in this discussion—beta-2 microglobulin, neopterin, and soluble interleukin-2 receptors—are relatively easy to measure and show a smaller degree of variability from assay to assay. Beta-2 microglobulin (part of the major histocompatibility complex class I antigen), neopterin (a product of stimulated macrophages), and soluble interleukin-2 receptors are all markers of immune activation. Moreover, these markers are systemic measures and may reflect the whole pool of cells affected by the disease, including macrophages. Of the three, beta-2 microglobulin has been found to be the most powerful predictor of progression to AIDS, although elevation of any of the three markers has been shown to be prognostic. Beta-2 microglobulin and neopterin are both abnormally elevated in more than 90 percent of people with symptomatic disease and in more than 40 percent of those with asymptomatic infection. With AZT therapy, all three have shown significant decreases and trends toward normalization. Indeed, some data have indicated that a decrease in beta-2 microglobulin early in the course of AZT therapy predicted the longer term clinical outcome. However, therapeutic studies of these markers are limited, which precludes judgments on whether drugs that influence the marker also affect adverse outcomes. In addition, the nonspecificity of these markers (they are elevated in the presence of numerous other viral, bacterial, and protozoan infections, as well as HIV) and their uncertain relationship to disease pathogenesis are major drawbacks to their use as surrogate endpoints in efficacy trials. Their use in combination with other markers (e.g., CD4+ cells, p24 antigen, IgA) has been proposed, and early studies have shown some promise.\*

The markers considered to this point all have qualities that support their potential use as surrogate endpoints for drug trials.\*\* It is well to reiterate, however, that any laboratory marker (or clinical marker, for that matter) being considered for such use should reflect an important physiological phenomenon related to the pathogenesis of the disease. Given the current state

\* Fahey and coworkers recently evaluated three cellular markers and five serologic markers for their ability to predict progression to AIDS. They found that progression to AIDS was predicted most accurately by the level of CD4+ cells in combination with the serum level of either neopterin or beta-2 microglobulin (Fahey, J.L., J.M.G. Taylor, R. Detels, B. Hofmann, R. Melmed, P. Nishanian, and J. Giorgi, "The prognostic value of cellular and serologic markers in infection with human immunodeficiency virus type 1," *New England Journal of Medicine*, vol. 322, no. 3, pp. 166-172 [1990]).

\*\* Other interesting but relatively less investigated surrogate endpoints were also mentioned during the conference discussion

of knowledge about the pathogenesis of AIDS, it is impossible to state absolutely that CD4+ lymphocyte decline reflects both necessary and sufficient factors in such pathogenesis. The caveats noted by researchers concerning whether the actual effect of a drug on CD4+ lymphocyte counts are really related to the pathogenic mechanisms of HIV infection should not be dismissed out of hand. Yet what must also be taken into account is the definitive finding that CD4+ cells are necessary for survival. In addition, although any one individual may be an exception, for populations of HIV-infected persons studies have shown an unequivocal association between low levels of CD4+ cells and measurable deleterious events that are AIDS-defining phenomena and, by extrapolation, predict survival. Thus, if one considers phenomena that are associated directly with deleterious events and that are found in everyone who is infected with HIV, one is left with a virus, the markers for which were summarized in the discussion on virology, and the CD4+ cell count, which invariably declines in HIV-infected individuals.

It may well be that damage occurs to the precursors of CD4+ cells during the initial precipitous drop in cell count following seroconversion and that this damage makes the loss of CD4+ cells inevitable and irreversible. If this were true, a steady, sustained increase in CD4+ cells might be too much to ask of any therapy. Perhaps, rather, the prevention of a decrease in such cells would be a more reliable indicator of the blunting of the effects of HIV. Many issues involving this marker remain to be resolved; nevertheless, with regard to the immunological parameters under discussion as surrogate endpoints, the CD4+ cell count must certainly stand out as the primary measurement of significance. Indeed, in some respects, patients have already accepted the CD4+ cell count as the major determination by which they judge drug efficacy in trials and decide when to end their participation in tests of what they believe are nonefficacious treatments. These very pragmatic judgments on CD4+ cell counts may bear increasingly on any policy regarding its use as a surrogate endpoint in efficacy trials.

#### **SPECIAL ISSUES IN THE MATERNAL AND PEDIATRIC POPULATION\***

AIDS in children and AIDS in adults differ in several respects, and these differences may affect the use of surrogate endpoints in general and the choice of any one endpoint in particular. For example, in the groups of infants and young children who have been identified as infected with HIV, the incubation period to AIDS—and, indeed, the whole course of the disease—is shorter than for adults: the median incubation in these groups of infants with either perinatally acquired or transfusion-associated AIDS is about two years. In addition, B lymphocytes in children appear to be more affected by the virus than are B cells in adults. The two groups are also susceptible to different types of opportunistic infections: severe bacterial infections are more common in childhood, although PCP and thrush, the most common opportunistic infections in children, are common to both groups. These factors affect the design and conduct of clinical trials, as does the fact that the pediatric AIDS population is substantially smaller than the population of adults with AIDS. Yet perhaps the most important difference between children and adults with AIDS is that infants and children are still developing. Thus, changes in the clinical markers of physical and mental development in children may have promise as early indicators of the action and efficacy of therapeutic drugs.

A number of natural history milestones are thought to have some prognostic value and may be useful to consider in the search for surrogate markers of drug effect in the pediatric population.

---

periods. One such candidate is delayed type hypersensitivity skin testing, particularly Candida skin testing, as a method for assessing immunocompetence in HIV-infected patients.

\*This section is based on material presented by Stanley Plotkin, Gwendolyn Scott, Catherine Wilfert, Karina Butler, Edward Connor, John Modlin, and Paul Parkman.

These factors can be divided into three types: immunological, virological, and clinical. The immunological and virological aspects of AIDS in children have much in common with the disease in adults; however, as noted above, the clinical picture in children varies substantially.

Of the immunological factors that seem to have prognostic value in children with AIDS, immunoglobulin levels appear to be one of the most significant. Most children with HIV disease have significantly elevated levels of immunoglobulin that may occur as early as three months of age and persist for long periods. Anecdotal information suggests that, late in the course of illness, the decline of these elevated immunoglobulin levels is associated with clinical deterioration in untreated children. Also, the small minority of children found to be hypogammaglobulinemic early in the course of disease seems to have a worse prognosis. However, as shown in the Phase I AZT trials, with therapy the majority of children in the trials had prompt and dramatic declines in measurable IgG and IgM toward normal levels. Thus, a decline in immunoglobulin levels associated with clinical improvement appears to be a marker of change when children are receiving drug therapy.

The prognostic relevance of CD4+ cell counts in children is more ambiguous. Lymphocyte counts in all children are higher than those of adults in the first year of life, and normal subset values have not been defined. At birth or during the first year, infected children who subsequently develop AIDS may have CD4+ cell counts of greater than 1,000. There may also be a significant increase in CD8+ cells, which produces the characteristic reversal of the CD4+/CD8+ ratio that in adults foreshadows the progression of clinical disease.

Several studies have shown drug effects on CD4+ cell counts in pediatric trials. Data from the Phase II pediatric AZT trials showed an initial increase in CD4+ counts with therapy. Preliminary data from the National Cancer Institute study that investigated a regimen of alternating ddC and AZT indicated that CD4+ cell counts increased and the CD4+/CD8+ ratio improved with therapy. In addition, preliminary results of a trial of ddI in children showed a similar improvement. Yet despite these promising results, the absence of data on normal CD4+ counts in children during the first year of life complicates the ability to determine abnormal levels. More data are needed on levels of CD4+ cells in children who are not HIV infected and in children with AIDS, over the course of the disease. Describing CD4+ cell values as percentages rather than as absolute numbers may also prove useful because, although children have approximately four times as many lymphocytes as adults, the percentage of CD4+ cells to total lymphocytes in children is comparable to that in adults.

In addition to CD4+ cell levels, several other immunological factors have been considered as laboratory markers in pediatric populations (and thus as potential surrogate endpoints for drug trials). These factors include tumor necrosis factor, beta-2 microglobulin, and soluble interleukin-2 receptors. Only limited data are available on the association of changes in these markers with disease progression; thus, no conclusions may be drawn regarding their value as surrogate endpoints in drug trials.

As in the case of immunological markers, virologic markers in children have not yet proved very useful for understanding prognosis and disease progression. Viral culture in children is used mainly for diagnosis rather than prognosis. Furthermore, as in adults, virus can be cultured from children on therapy. PCR studies in children (see the general discussion of this technique under the virological panel) are limited and have concentrated on the early neonatal period, once again, mainly for the purposes of diagnosis. This method needs further quantification and standardization before it is ready for use in drug evaluation.

Among children as among adults, the virologic marker that so far has stirred the most interest regarding its possible use as a surrogate endpoint of drug effect is the level of p24 antigen. In spinal fluid, p24 antigen has been recognized as an early indication of HIV infection. Other data relating to p24 antigen in children suggest a correlation between quantitative levels of antigen and disease stage. Studies in New York, New Jersey, and the Netherlands have found a correlation between the presence of high levels of antigen and disease progression. The p24 antigen has also shown changes in the presence of drug therapy, declining in response to AZT therapy in Phase I and II trials. The

NCI study of ddC plus AZT reported brisk falls in levels of the antigen. In addition, data from the collaborative AZT studies indicated that essentially all of the children who had detectable p24 antigen in spinal fluid at the onset of therapy had reductions of the antigen to undetectable levels by 24 weeks. These reports are promising; however, as in adults, a major drawback to the use of p24 antigen levels as a surrogate endpoint is that not all children with HIV infection have detectable antigenemia.

Certain clinical aspects of AIDS in children have been shown to have prognostic value in the course of untreated infection; however, their utility as surrogate markers of drug effect is unclear. Prognostic clinical factors include age at presentation of clinical disease, the type of disease with which the child presents, and the development of opportunistic infection. The particular type of opportunistic infection developed by a child may also be a key to the rate of disease progression. In general, children who present with AIDS at less than one year of age have a shorter median survival than those who present after one year. The type of clinical disease with which the child presents also appears to predict prognosis: for example, survival following PCP is dismal. In adults, the presence of PCP is associated with low CD4+ cell counts; in children, there is less correlation with the absolute number of CD4+ cells (children have died with CD4+ counts of 1,000 to 2,000). On the other hand, children who develop lymphoid interstitial pneumonitis (LIP) have a longer median survival. (Although children with LIP meet the CDC case definition for AIDS, they are actually less symptomatic and are sometimes combined with ARC patients in clinical trials.)

Finally, a unique aspect of children is that the "business" of infancy and childhood is growth and development. These processes in children are predictable and well-described; because they also proceed rapidly, changes can be detected early. Loss of appetite, wasting, and neuropsychological impairment (detectable in 40 to 95 percent of HIV-infected children) are hallmarks of HIV infection in this population. Consequently, measurements of linear growth and weight gain and assessments of neurobehavioral function through tests of motor skills and cognitive ability can be used to evaluate the action of a drug.

Several studies have shown what many researchers believe is evidence of drug efficacy using these developmental measures. Data from the multicenter Phase I trial of AZT in children showed that administration of the drug was associated with weight gain in a population of children that had shown an abnormal distribution of weights by percentile at their enrollment in the trial. The evidence of growth that occurred when the children were placed on therapy was an indication that their disease had undergone alteration. Changes were also reported following AZT therapy in tests of both gross and fine motor skills and cognitive function. Before therapy, the distribution of children in the trial by percentile on a scale measuring cognitive function was abnormal; following therapy, the distribution was much more in line with normal values. In a similar Phase I study of AZT at NCI, 13 of 21 children had neurological impairment on entering the trials, and all 13 showed neurologic improvement with AZT. However, this neurologic improvement appeared to be independent of other disease symptoms and was not always accompanied by improvement in immunologic function.

One important feature of these data is that, although some children on therapy may not have actually improved their status, very few children continued to show declines in development. Thus, the stability of neurobehavioral status and the stabilizing power of a pharmacological agent should be considered both in conclusions about efficacy and discussions of appropriate indicators of drug action. Many pediatric researchers see these types of developmental measures as some of the most promising candidate markers, although there are some drawbacks to neurodevelopmental tests. For example, they are more reliable in children who are older than two or three years, and they are extremely expensive to administer. In addition, interobserver variability may be a problem, particularly in multicenter trials.

Some pediatric researchers have called for the development of a comprehensive, well-designed clinical scoring system as a potential outcome measure for drug effect. Such a system could incorporate growth, neurological development, development of opportunistic infection, and dysfunction

of particular organ systems (e.g., heart, lungs, liver, and kidneys). There is precedent for employing a clinical scoring system that reflects overall well-being in evaluations of interventions for patients with cystic fibrosis.

Any discussion of treatment for pediatric AIDS would be incomplete without considering prevention or interruption of perinatal transmission using chemotherapeutic modalities. Vertical transmission of HIV from infected mothers to their infants accounts for 80 percent of all childhood cases of AIDS. Studies are currently planned to investigate chemoprophylaxis against in utero transmission (i.e., prophylaxis of the fetus) and intrapartum transmission (prophylaxis of the newborn). Clinical endpoints of drug efficacy in such studies would be the presence or absence of established infection in the infants.

With the apparent short incubation period of pediatric HIV infection (in those children observed thus far) and a generally short disease course characterized by early clinical signs and symptoms, the question arises: Is there a need for surrogate markers to speed definitive drug trials in children? The consensus of the conference was that surrogate markers are important for trials in children because analysis of early trends in responses to experimental drugs can identify drugs that merit further evaluation and direct future research. Moreover, as in adults, the hope is that new drugs can eventually be approved sooner on the basis of surrogate endpoints.

Indeed, the speed and sequencing of pediatric (versus adult) drug trials have become issues of some debate with the advent of AIDS, and several points are worth noting. More than two years have passed since AZT was approved for adults; yet at the time of the conference it had not received full approval for marketing in children. The conditions that promote such a substantial, and, as many conferees expressed it, unacceptable lag in drug availability for children are forcing a reconsideration of the drug approval process for this population. In the past, sponsors and researchers have waited for evidence of efficacy from adult trials before commencing trials in children. Now, however, many in the field agree this should no longer be the case; rather, Phase I studies for safety, tolerance, and pharmacokinetics, as well as preliminary evidence of efficacy, generally should proceed in parallel in adults and children. (In some instances, preliminary studies to establish general parameters of dosage and safety in a small number of adults may be useful, particularly in new classes of drugs and to serve as a basis of selection for Phase I trials in children among a number of candidate drugs.) All drugs that will be prescribed for children should be tested in children for safety and pharmacokinetics (Phase I trials). On the other hand, some efficacy trials may be easier to conduct in adults, by virtue of the larger number of adult patients, and these efficacy trials (Phases II or III) should not necessarily require duplication in children before a drug is approved for pediatric use. For example, data from adult studies on the efficacy of AZT in infected persons with CD4+ counts of less than 500 have been applied to children to support the practice of treating symptomatic children and infected children with CD4+ counts of less than 500 with AZT. (Future studies may reveal that a higher CD4+ count would be a more appropriate cutoff in infants.) Furthermore, there may, in the future, be drugs particular to pediatric AIDS that may never undergo evaluation in adults. Randomized, controlled clinical trials will continue to be essential in most cases in children and can be efficiently conducted because of the apparent short incubation period and rapid progression of perinatally acquired infection. Trials using placebo are ethical at this time only for newborn infants of infected mothers.

#### ENDPOINTS, TRIAL DESIGN, AND REGULATION\*

Using surrogate endpoints for drug efficacy trials raises the additional problem of how to design such trials to show the effects of a drug in an era of combination chemotherapy, prophylaxis for some

---

\* This section is based on material presented by Frank Young, Ellen Cooper, Martin Delaney, Sandra Lehrman, Daniel Hoth, Peter Barton Hutt, and Stephen Lagakos.

opportunistic infections, and an established standard (AZT). The development of AZT has ended the days of placebo-controlled trials for antiretroviral agents in symptomatic patients. The urgency of the epidemic, and the recent appearance in some patients of virus with in vitro resistance to AZT, has brought increasing pressure for quick approval of any drug that appears to proffer benefits. This force in turn has created a dynamic tension between the traditionally measured response of regulators, who are responsible for careful review of the data that seek to establish drug safety and efficacy, and patients and their advocates, whose overriding concern is to provide promising therapies to infected persons as quickly as possible. Balancing the requirements of both these groups in the search for ways to speed drug approval has led to the recent focus on the potential use of surrogate endpoints in clinical trials.

The current regulatory position is that no single marker has been sufficiently well validated to be used as a surrogate endpoint for drug efficacy trials. Rather, FDA sees the use of combinations of markers, including intermediate clinical endpoints (less severe or life-threatening clinical illness), as the direction it would prefer continuing studies to take. Its present policy is to encourage the investigation of these markers within clinical trials as a way to gather the additional data needed to assess their validity as endpoints. One of the factors working against any quick decision on the use of such endpoints is the chronic aspect of HIV infection; the safety and efficacy of drugs must be assessed over the length of what is now known to be a long disease course. Consequently, it is virtually impossible to provide the quick answers desired by those who have the disease or who must treat patients daily.

Technically, there is nothing in any of the regulatory statutes governing the drug approval process that precludes FDA's approving a drug on the basis of its effect on a surrogate endpoint. From the FDA's point of view, however, the establishment of such an endpoint or endpoints is premature at this time. Further progress in the following areas is needed before surrogate endpoints can be established: the ability to detect and measure markers reliably, the establishment of standards for significant changes in markers, the possibility of using different markers at different phases of drug development or with different classes of compounds, and, most important, the correlation of changes in these markers with both short- and long-term clinical outcome. To address these issues requires the prospective collection of more data that consistently establish the relationship between changes in clinical condition and changes in surrogate markers over time at different stages of disease and with different classes of drugs. In such an effort, proper statistical analysis is critical, especially in light of the great variability observed in all of the laboratory markers currently under consideration.

AIDS patients and advocates argue that the use of surrogate endpoints will speed the availability of drugs and in clinical studies may provide a type of "safety net" for patients by defining a threshold below or beyond which patients should not be allowed to progress without the opportunity to receive therapeutic interventions. In particular, AIDS patient advocates argue for the acceptance of measures of intermediate clinical endpoints as surrogates, including the so-called quality-of-life indicators. Some have proposed using such indicators for a form of conditional licensing for a promising therapy whose effectiveness could then be confirmed (or disproved) in postmarketing surveillance. In addition, they suggest the possibility of revisiting and reviewing past studies and data in the light of new knowledge to ensure that no promising drug has been overlooked or rejected on the basis of what is now known to be inaccurate or incomplete information.

FDA officials have stated that the agency would approve a drug that had been shown to have a clearly beneficial effect on intermediate clinical variables, even though the drug had not been shown to affect the virus. The issue turns, however, on how definitively clinical benefit has been demonstrated. The results of small uncontrolled trials to measure improvement in these clinical phenomena could well be confounded by operation of the placebo effect (patients improve because they think they are receiving an effective therapy). As a way to facilitate decisions regarding surrogate endpoints for drug trials, FDA officials expressed the need for a public consensual process. Most FDA decisions are fundamentally driven by the current standard of opinion in the

medical community about a particular issue. Once this standard has been more clearly defined and enunciated—that is, once the data are available from which to develop such a standard—the use of surrogate endpoints in drug efficacy trials may be more widely implemented and carry more weight in judging the effectiveness of a therapy against HIV and AIDS.

As in many aspects of HIV infection and AIDS, the lack of data on the mechanisms of action of the laboratory markers restricts both the ability of researchers to reach conclusions on the advantages of any particular marker and the ability of regulators to help speed drugs to the marketplace and into the hands of patients. As one way to address the data problem the ACTG has now moved to more multicenter trials of drugs and dosages. Still, at least for the present, there will continue to be a number of factors that are not well understood but that may have a substantial effect on whether and how a particular surrogate marker is used. For example, HIV infection is not a single entity—it manifests differently in different individuals and at different stages of disease. There is also some evidence that demographic factors may affect prognosis, an aspect of the problem that continues to surface but that has not as yet been substantially investigated. Furthermore, as in any new area of medicine for which remedies are being sought, the question of the individual patient looms large. The available studies of markers deal, as they must, with populations; the use of markers for clinical staging of individual patients requires further study.

Decisions regarding the use of any of the markers discussed at the conference rest on determining what level of risk is tolerable and at what point the medical and regulatory communities should stop seeking perfection in the state of knowledge and make decisions based on the data that are available. There can be no doubt that what might be called the classic, large-scale controlled trial produces the best information on the efficacy of a drug; without such a trial, AZT would not have been approved as quickly nor could it be prescribed for AIDS as confidently. The heavy demands of the disease, however, argue for thoughtful consideration of less traditional methods that may speed drug approval and make effective therapies available to patients. Surrogate endpoints merit study and development as one such approach.

#### ROUNDTABLE SUMMARY\*

In the climate of urgency that surrounds the search for efficacious therapies against a deadly disease, surrogate markers of drug effectiveness have great appeal as a mechanism for speeding determinations of efficacy and safety. Their utility, however, is not limited to the arena of drug regulation. Clinical management of patients, for example, will also be influenced by the information collected about surrogate endpoints. Indeed, CD4+ cell counts are now widely used to evaluate individual patients' disease status and their suitability for medical interventions. Choosing among the candidate markers that have been identified as possibly useful thus assumes great importance, especially considering that those surrogates identified today by researchers as most promising will influence future clinical trial design. The costs of clinical trials and the need to separate less promising endpoints from those that are most closely related to the pathogenesis of the disease and correlate most strongly with the true clinical endpoints of morbidity and mortality argue for careful consideration of the endpoints that may currently be considered candidate surrogates.

Of the laboratory markers of disease progression that have been considered so far, the CD4+ cell count is preeminent. Drugs may have two types of beneficial effect on CD4+ cell counts. In patients with low CD4+ cell counts, they may produce a rise in CD4+ cells. In patients whose CD4+ cells indicate that they are not yet at high risk for deleterious events, beneficial drugs may decrease the rate of decline in CD4+ cell counts. For that period of time during which these patients have higher CD4+ counts than they otherwise would have had, their risk of adverse outcomes is

---

\*This section is a synthesis of the information presented at the conference based on the Roundtable members' discussion of the material immediately after the conference concluded.

presumably diminished. Although the data on the relationship between CD4+ cell count and prognosis in the natural history of HIV infection are much stronger than those for the significance of CD4+ changes in studies of drug effect (in which analyses are still preliminary), the early evidence is promising. In addition, the clear relationship of CD4+ cell counts to disease pathogenesis adds weight to this conclusion. Data collected and analyzed in the future, of course, may modify this assessment of the marker's promise. In addition, as is the case for all surrogate endpoints, clinical measurements of significance must always take precedence in assessing the effectiveness of a drug. Nevertheless, with this caveat in mind, and all other things being equal, the available evidence suggests that any drug that raises CD4+ cell counts significantly or delays their decline to a level predictive of adverse events should be considered effective for that measure, and therefore for the disease. This finding should be an important consideration in the drug approval process for AIDS therapeutics.

Other markers of immune activation—neopterin, beta-2 microglobulin, and soluble interleukin-2 receptors—are also of interest. These markers correlate with disease activity and deserve further study to establish how they perform in therapeutic studies. At this time, however, their relationship to the pathogenesis of HIV disease is less clear than that of CD4+ cell counts. Thus, their best use may be in combination with CD4+ cell counts, percentages, or CD4+/CD8+ ratios.

Because there is substantial variation in viral burden during the course of disease, measuring the virus itself may be the most direct link to disease pathogenesis and offers promise as a surrogate endpoint of drug effect. Until recently, viral cultures have not appeared to be useful surrogates because they gave only a yes/no answer to the question of whether virus was actually present (current techniques are sensitive enough to allow investigators to culture virus from peripheral blood mononuclear cells at all stages of disease) and could not be used to quantify viral burden. Now, however, promising early results have appeared showing that the magnitude of plasma viremia correlates well with disease stage. In addition, some researchers have now demonstrated that plasma viremia titers decline in response to AZT therapy.

As a measure that reaches to the heart of disease pathogenesis, quantification of viral burden offers clear promise as a marker of disease progression and a possible surrogate endpoint of drug effect. Further development of good quantitative viral assays should be encouraged. The technique of quantitative polymerase chain reaction (PCR) also shows substantial potential for quantitating viral burden. However, its early stage of development and lack of standardization preclude its immediate use as a definitive measure of drug effect.

Another virological marker investigated for its potential as a surrogate endpoint is p24 antigenemia. Although levels of p24 antigen do seem to correlate with drug effect, there are some important limitations to their usefulness. For one thing, only a small proportion of HIV-infected persons has measurable p24 antigen levels. Therefore, using p24 as the major indicator of drug effect would severely restrict the population of patients eligible for drug trials. On the other hand, measurement of p24 antigen is relatively inexpensive and easy to perform, and in patients in whom p24 is present it does correlate with drug effect. For these reasons, further information on the response of p24 antigen levels to therapy will be valuable to collect.

Less severe or life-threatening clinical illness may offer another realm of opportunity in the search for appropriate measures in drug trials. An essential distinction between clinical and laboratory endpoints is that, although less severe clinical endpoints may be considered surrogates insofar as they correlate with disease progression and regression, they may have important significance for patient well-being by themselves. Therefore, any drug that has beneficial effects on an intermediate clinical marker may be important in two discrete ways: for its effect on the clinical condition as an end in itself and for its effect on further disease progression. For example, a drug that effectively delays the onset of bacterial infections would be important for those effects as well as any effect it might have on delaying overall disease progression. For this reason, some people object to labeling any clinical measure a surrogate endpoint. The natural history of HIV infection indicates that earlier clinical illness is detectable and that such illness itself predicts more severe subsequent



clinical outcomes. For example, tuberculosis, particular bacterial infections such as pneumonia, skin infections, diarrhea, fever, weight loss, minor cognitive changes, changes in Karnofsky scores, and objective measures of quality-of-life outcomes demand further study for their relation to disease progression and for their potential role as early indicators of drug effect.

Of necessity, the conference purview was limited mainly to consideration of laboratory surrogate markers. In the pediatric population, however, persuasive evidence that clinical endpoints may be the most powerful predictors of drug effect in infants and children overrode that limitation. Pediatricians are accustomed to measuring growth and development; thus, the capability for accurate quantification is already in place. In addition, studies have shown that the well-documented failure to grow normally that is found in infants and young children with HIV infection is clearly altered with antiretroviral therapy. Researchers in this area take this phenomenon to mean that antiretroviral therapy has had an effect on overall disease progression. Reverses in neurobehavioral abnormalities, which are more specific to HIV infection than are growth abnormalities, also provide objective, measurable evidence that antiretroviral therapy has affected disease progression.

Most studies of the relations between laboratory markers and disease progression and regression have used adult populations. Consequently, although certain trends in such relations are comparable from adults to children, the predominant theme here is the paucity of data from sequential studies in infected children. One problem in understanding the importance of CD4+ cell counts for pediatric patients is the lack of established norms in infancy and early childhood. Because normal levels of these cells are higher in infants than in adults, the meaning of exact numbers of CD4+ cells and their relation to disease progression (for example, the often-used level of 500 CD4+ cells) may differ in these populations. One potential solution to this problem may be for researchers to adopt the convention of expressing CD4+ cells as a percentage of total lymphocytes rather than as an absolute value.

In some respects, the exigencies of AIDS and HIV infection have produced a change in thinking about clinical trials, especially in the case of pediatric populations. Fueling such a reconsideration of traditional methods and established practice has been the realization that those mechanisms designed to protect infants and young children from the potentially damaging effects of drugs have paradoxically retarded the delivery of effective therapy to this population. For example, in the past, trials in children with any particular agent were not initiated until the drug had been proved effective through large-scale efficacy trials in adults. Now, however, a consensus is forming that the initiation of clinical trials in infants and children with HIV infection and approval of drugs for them should not await action in the adult population but may proceed in parallel. Because toxicity of drugs may differ in children and adults, Phase I pediatric trials to establish safety are always necessary; however, when evidence from large-scale adult trials shows the effectiveness of a drug, large-scale Phase II and III studies for efficacy do not necessarily require duplication in children. In other cases, randomized controlled clinical trials in children will continue to be practical and important, especially in those cases in which a particular drug is only being considered for use in the pediatric age group.

The search for surrogate endpoints of drug efficacy as a way to shorten clinical trials of drug safety and effectiveness is impelled by the understandable urgency that attends a disease such as AIDS. On the basis of currently available evidence, CD4+ cell counts appear to be the marker most immediately applicable to the drug approval process. In addition, plasma viremia holds great promise. It is now essential to pursue the clinical studies that will clarify the value of these and other potential surrogate markers and their relationship to ultimate clinical outcome by incorporating their analyses into drug trial design.